

## CLAIMS

That which is claimed:

1. A method of detecting the presence of an installation on a data processing system, comprising:
  - providing a signature that comprises  $m$  files having paths associated therewith, respectively;
  - 5 determining a number  $n$  files on the data processing system that match files in the signature;
  - determining a files found ratio given by  $n/m$ ;
  - applying a transformation to the signature by replacing at least a portion of at least one of the paths with a new path; then
  - 10 determining a distance between the  $n$  files on the data processing system and the  $m$  signature files, the distance corresponding to a sum of a number of path segments associated with the  $m$  signature files that cannot be matched to a corresponding path segment associated with files on the data processing system; and
  - determining if the installation is present on the data processing system based on the files found ratio and the distance.
  - 15
2. The method of Claim 1, further comprising:
  - determining if the files found ratio is greater than a files found threshold; and
  - wherein determining the distance and determining if the installation is present
  - 20 are performed if the files found ratio is greater than the files found threshold.
3. The method of Claim 1, wherein applying the transformation comprises:
  - selecting a transformation based on a path associated with one of the  $m$
  - 25 signature files; then
  - determining the distance between the  $n$  files on the data processing system and the  $m$  signature files using the selected transformation;
  - repeating selecting the transformation and determining the distance between the  $n$  files on the data processing system and the  $m$  signature files using the selected
  - 30 transformation for each one of the  $m$  signature files; and

applying a transformation to the signature that is associated with the smallest distance.

4. The method of Claim 1, wherein if one of the  $m$  signature files matches  
5 a plurality of files on the data processing system, then the distance between the one of the  $m$  signature files and the plurality of matching files on the data processing system is a smallest distance between the one of the  $m$  signatures files and any one of the plurality of matching files on the data processing system.

10 5. The method of Claim 1, wherein determining if the installation is present comprises:  
comparing the files found ratio to a files found threshold;  
comparing the distance to a distance threshold; and  
determining that the installation is present on the data processing system if the  
15 files found ratio exceeds the files found threshold and the distance is less than the distance threshold.

6. The method of Claim 5, further comprising:  
selecting the files found threshold based on the number of signature files  $m$   
20 and/or whether the installation has at least one file associated therewith whose installation is optional.

7. The method of Claim 1, wherein determining the number  $n$  files on the data processing system that match files in the signature comprises:  
25 searching the data processing system to find a first file that matches a first one of the  $m$  signature files, the found file having a distance  $d_{max}$  associated therewith;  
searching the data processing system to determine if another file that matches another one of the  $m$  signature files can be found within the distance  $d_{max}$ ;  
ending the search of the data processing system if the other file that matches  
30 the other one of the  $m$  signature files is found within the distance  $d_{max}$ ;  
updating the distance  $d_{max}$  with a value of a distance associated with the other file that matches the other one of the  $m$  signature files if the other file is not found within the distance  $d_{max}$ ; and

performing searching the data processing system to determine if another file matches another one of the  $m$  signature files, ending the search, and updating the distance  $d_{max}$  for each of the  $m$  signature files.

5           8.       The method of Claim 1, wherein the installation comprises a software program, a firmware program, a documentation file, a printer font, and/or a web file.

          9.       A system for detecting the presence of an installation on a data processing system, comprising:

10           means for providing a signature that comprises  $m$  files having paths associated therewith, respectively;

          means for determining a number  $n$  files on the data processing system that match files in the signature;

          means for determining a files found ratio given by  $n/m$ ;

15           means for applying a transformation to the signature by replacing at least a portion of at least one of the paths with a new path;

          means for determining a distance between the  $n$  files on the data processing system and the  $m$  signature files, the distance corresponding to a sum of a number of path segments associated with the  $m$  signature files that cannot be matched to a

20           corresponding path segment associated with files on the data processing system; and

          means for determining if the installation is present on the data processing system based on the files found ratio and the distance.

          10.      The system of Claim 9, further comprising:

25           means for determining if the files found ratio is greater than a files found threshold;

          wherein the means for determining the distance comprises means for determining the distance between the  $n$  files on the data processing system and the  $m$  signature files if the files found ratio is greater than the files found threshold; and

30           wherein the means for determining if the installation is present comprises means for determining if the installation is present on the data processing system based on the files found ratio and the distance if the files found ratio is greater than the files found threshold.

11. The system of Claim 9, wherein the means for determining if the installation is present comprises:

- means for comparing the files found ratio to a files found threshold;
- 5 means for comparing the distance to a distance threshold; and
- means for determining that the installation is present on the data processing system if the files found ratio exceeds the files found threshold and the distance is less than the distance threshold.

10 12. The system of Claim 11, further comprising:

means for selecting the files found threshold based on the number of signature files  $m$  and/or whether the installation has at least one file associated therewith whose installation is optional.

15 13. The system of Claim 9, wherein the means for determining the number  $n$  files on the data processing system that match files in the signature comprises:

- means for searching the data processing system to find a first file that matches a first one of the  $m$  signature files, the found file having a distance  $d_{max}$  associated therewith;
- 20 means for searching the data processing system to determine if another file that matches another one of the  $m$  signature files can be found within the distance  $d_{max}$ ;
- means for ending the search of the data processing system if the other file that matches the other one of the  $m$  signature files is found within the distance  $d_{max}$ ; and
- means for updating the distance  $d_{max}$  with a value of a distance associated
- 25 with the other file that matches the other one of the  $m$  signature files if the other file is not found within the distance  $d_{max}$ .

14. The system of Claim 9, wherein the installation comprises a software program, a firmware program, a documentation file, a printer font, and/or a web file.

30

15. A computer program product for detecting the presence of an installation on a data processing system, comprising:

a computer readable storage medium having computer readable program code embodied therein, the computer readable program code comprising:

computer readable program code configured to provide a signature that comprises  $m$  files having paths associated therewith, respectively;

5 computer readable program code configured to determine a number  $n$  files on the data processing system that match files in the signature;

computer readable program code configured to determine a files found ratio given by  $n/m$ ;

10 computer readable program code configured to apply a transformation to the signature by replacing at least a portion of at least one of the paths with a new path;

computer readable program code configured to determine a distance between the  $n$  files on the data processing system and the  $m$  signature files, the distance corresponding to a sum of a number of path segments associated with the  $m$  signature files that cannot be matched to a corresponding path segment associated with files on  
15 the data processing system; and

computer readable program code configured to determine if the installation is present on the data processing system based on the files found ratio and the distance.

16. The computer program product of Claim 15, further comprising:

20 computer readable program code configured to determine if the files found ratio is greater than a files found threshold;

wherein the computer readable program code configured to determine the distance comprises computer readable program code configured to determine the distance between the  $n$  files on the data processing system and the  $m$  signature files if  
25 the files found ratio is greater than the files found threshold; and

wherein the computer readable program code configured to determine if the installation is present on the data processing system comprises computer readable program code configured to determine if the installation is present on the data processing system based on the files found ratio and the distance if the files found  
30 ratio is greater than the files found threshold.

17. The computer program product of Claim 15, wherein the computer readable program code configured to determine if the installation is present comprises:

computer readable program code configured to compare the files found ratio to a files found threshold;

computer readable program code configured to compare the distance to a distance threshold; and

5 computer readable program code configured to determine that the installation is present on the data processing system if the files found ratio exceeds the files found threshold and the distance is less than the distance threshold.

18. The computer program product of Claim 17, further comprising:

10 computer readable program code configured to select the files found threshold based on the number of signature files  $m$  and/or whether the installation has at least one file associated therewith whose installation is optional.

19. The computer program product of Claim 15, wherein the computer  
15 readable program code configured to determine the number  $n$  files on the data processing system that match files in the signature comprises:

computer readable program code configured to search the data processing system to find a first file that matches a first one of the  $m$  signature files, the found file having a distance  $d_{max}$  associated therewith;

20 computer readable program code configured to search the data processing system to determine if another file that matches another one of the  $m$  signature files can be found within the distance  $d_{max}$ ;

computer readable program code configured to end the search of the data processing system if the other file that matches the other one of the  $m$  signature files is  
25 found within the distance  $d_{max}$ ; and

computer readable program code configured to update the distance  $d_{max}$  with a value of a distance associated with the other file that matches the other one of the  $m$  signature files if the other file is not found within the distance  $d_{max}$ .

30 20. The computer program product of Claim 15, wherein the installation comprises a software program, a firmware program, a documentation file, a printer font, and/or a web file.